

基于 YOLOX-NGS 的群养猪只攻击行为识别

李艳文, 李菊霞^{*}, 纳腾潇, 智晴宇, 段磊, 张朋鹏

(山西农业大学信息科学与工程学院, 太谷 030801)

摘要:为解决复杂猪舍环境下猪只堆叠和粘连导致群养猪只攻击行为识别准确率低和有效性差的问题, 该研究提出一种改进的 YOLOX 模型, 引入攻击活动比例 (P_{AA}) 和攻击行为比例 (P_{AB}) 2 个优化指标, 对群养猪只的撞击、咬耳和咬尾等典型攻击行为进行识别。首先, 为提高模型特征提取能力添加归一化注意力模块获取 YOLOX 颈部的全局信息; 其次, 将 YOLOX 中的 IoU 损失函数替换为 GIoU 损失函数, 以提升识别精度; 最后, 为保证模型的实时性将空间金字塔池化结构 SPP 轻量化为 SPPF, 增强检测效率。试验结果表明, 改进的 YOLOX 模型平均精度达 97.57%, 比 YOLOX 模型提高 6.80 个百分点。此外, 当 P_{AA} 和 P_{AB} 阈值分别为 0.2 和 0.4 时, 识别准确率达 98.55%, 有效解决因猪只攻击行为动作连续导致单帧图像行为识别可信度低的问题。研究结果表明, 改进的 YOLOX 模型融合 P_{AA} 和 P_{AB} 能够实现高精度的猪只攻击行为识别, 为群养生猪智能化监测提供有效参考和技术支持。

关键词: 猪; 行为识别; 检测; 攻击行为; 注意力模块; 损失函数; 特征金字塔

doi: 10.11975/j.issn.1002-6819.202308132

中图分类号: TP391.41; TP183

文献标志码: A

文章编号: 1002-6819(2023)-24-0177-08

李艳文, 李菊霞, 纳腾潇, 等. 基于 YOLOX-NGS 的群养猪只攻击行为识别[J]. 农业工程学报, 2023, 39(24): 177-184.

doi: 10.11975/j.issn.1002-6819.202308132 <http://www.tcsae.org>

LI Yanwen, LI Juxia, NA Tengxiao, et al. Recognizing attack behavior of herd pigs using improved YOLOX[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(24): 177-184. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202308132 <http://www.tcsae.org>

0 引言

在猪只养殖过程中, 猪只攻击行为的识别主要以人工监测为主, 不仅管理成本高, 同时可能引起猪只的应激反应, 存在人畜交叉感染的风险^[1]。随着计算机视觉技术的发展, 利用该技术对群养猪只进行非接触式、低应激的健康监测成为可能, 猪只目标检测和行为识别是实现猪只自动化监测与智能化分析的前提^[2]。快速、准确的猪只攻击行为检测, 对实现精准化、个性化的猪只健康智能养殖具有重要意义。

目前, 国内外已有研究者针对猪只攻击行为的检测识别做出相应研究。VIAZZI 等^[3]针对猪只运动图像的历史数据提取运动特征, 采用线性判别分析对猪只异常行为进行分类, 准确率达到 89.0%。OCAZK 等^[4]通过统计猪只活动指数, 采用 BP 神经网络对猪只攻击行为的程度进行分类, 对攻击行为分类准确率达 99.2%。JONGUK 等^[5]通过 Kinect 深度传感器提取猪只活动特征, 采用支持向量机对猪只追逐和攻击两种异常行为进行分类, 分类准确率达到 90.2%。CHEN 等^[6]通过分析相邻关键帧图像之间猪只目标的位移变化, 从而提取猪只加速度特

征, 采用分层聚类方法对猪只攻击行为程度进行分类, 准确率达到 95.82%。以上对猪只攻击行为检测的研究大都是基于图像处理技术提取图像中的某种特征, 有较高的准确率, 但需要进行大量的人工预处理, 再结合机器学习方法进行处理。在实际应用中, 这些方法存在效率低、工作量大等问题。随着深度学习在其他领域中展现出强大的学习泛化能力, 且在猪只行为检测领域也得到大量的应用^[7-9]。YANG 等^[10]基于 Fast R-CNN 检测猪只进食行为; 使用全卷积神经网络的方法分割母猪和仔猪, 使用母猪几何特征提取对应的空间信息以识别母猪哺乳行为^[11]。杨秋妹等^[12]使用卷积神经网络针对个体猪只饮水行为做出相应的研究。李菊霞等^[13]采用 YOLOv4 模型对猪只饮水和进食行为进行识别。高云等^[14]采用深度学习的方法, 搭建 3D 卷积神经网络模型, 将其应用于群养猪猪侵略性行为的识别。相较于传统的机器学习方法, 深度学习通过对低维特征到高维特征的提取学习, 能够做到绝大部分场景下各类任务的检测识别^[15-17]。深度学习在猪只单一行为的研究中展现出优异的性能, 但针对群养猪只的交互行为研究较少。此外, 在实际养殖场景中, 猪只行为特征提取和识别以单帧图像为基础, 猪只攻击行为识别可信度低, 需要对视频连续帧做进一步分析。

针对上述问题, 本试验采用帧间差分法^[18]获取关键帧构建猪只攻击行为数据集, 通过改进 YOLOX 模型提高猪只攻击行为的特征提取能力和检测精度; 引入攻击活动比例 (P_{AA}) 和攻击行为比例 (P_{AB}) 两个优化指标

收稿日期: 2023-08-16 修订日期: 2023-11-22

基金项目: 山西省基础研究青年项目 (202103021223141)

作者简介: 李艳文, 硕士, 讲师, 研究方向为机器视觉与农业图像处理。

Email: lyw@sxau.edu.cn

※通信作者: 李菊霞, 硕士, 教授, 研究方向为农业信息化。

Email: lijxsn@126.com

对连续帧进一步量化, 判定猪只攻击行为, 以有效提高猪只攻击行为的识别效果。

1 数据采集与预处理

1.1 数据采集

本试验数据于 2020 年 7 月在山西省临汾市汾西县猪只培育基地采集。选取 9 只 5 月龄育肥猪, 饲养在规模为 4.5 m×4.0 m×2.8 m 的封闭猪舍内。试验使用海康威视 DS-2CD3345D-I 型号摄像头, 在自然光下采取向下倾斜 60°视角采集数据。该角度与平视和俯视相比, 可以获取更丰富的猪只行为特征, 避免出现大面积遮挡。

1.2 猪只攻击行为定义

本试验数据猪只的攻击行为是群养猪只中多头猪只的状态交互, 是一个复杂、渐进的过程。在发生攻击行为初期, 猪只通过鼻子嗅闻、轻推等方式进行初步试探, 随后攻击行为逐渐加剧, 往往伴随着更激烈的挤压、咬、撞击等行为。在攻击行为最剧烈的时候, 会出现咬耳朵、咬尾巴和咬身体等行为。撕咬通常持续时间长, 往往会造成皮肤损伤、创口等。本试验中定义的攻击行为是撕咬、撞击, 如表 1 所示。

表 1 攻击行为定义

Table 1 Definition of aggressive behavior

攻击种类 Type of aggressive	攻击行为名称 Name of aggressive behavior	行为描述 Behavior description
撕咬 Bite	咬耳朵	撕咬另外一只猪只的耳朵
	咬尾	撕咬另外一只猪只的尾巴
	咬身体	撕咬另外一只猪只的身体部位, 包含嘴部, 颈部
撞击 Knock	头对头撞击	猪只用头部撞击另一只猪的头部
	头对身撞击	猪只用头部撞击另一只猪的身体部位

在数据采集过程中, 对猪只日常行为视频进行反复的观察, 截取出 185 段具有攻击行为的猪只视频片段; 由于猪只运动缓慢、休息时间较长, 使用连续帧作为训练集时, 训练过程中出现大量的重复样本, 导致模型拟合不足。为提高模型的拟合优度, 本试验采用帧间差分法提取关键帧作为有效图像。当两帧之间像素灰度值差值的总和大于预先设定的阈值时, 将第二帧定义为关键帧。关键帧判别函数为:

$$D(i, j) = \begin{cases} 1, & \sum |f_i(x, y) - f_j(x, y)| > T_i (j > i) \\ 0, & \text{other} \end{cases} \quad (1)$$

式中 $f_i(x, y)$ 表示第 i 帧图像所对应的 (x, y) 像素点的灰度值, $f_j(x, y)$ 表示第 j 帧图像所对应的像素点的灰度值。 T_i 为两帧之间像素灰度值差值的总和, 如果帧间差值的绝对值大于阈值 T_i , 则 $D(i, j)=1$, 表示第 j 帧是相对于第 i 帧的关键帧, 否则 $D(i, j)=0$, 表示第 j 帧不是关键帧。

在获得关键帧后, 使用以下 2 个步骤对获得的图像进行处理, 得到最终的试验数据集。

1) 将输入图像的像素分辨率从 1920×1080 调整到模型训练所需的 640×640。

2) 使试验所需的数据集符合 PASCALVOC 格式要

求, 并对数据集进行标注, 将图像中有攻击行为的猪只标记为攻击。如图 1 所示。

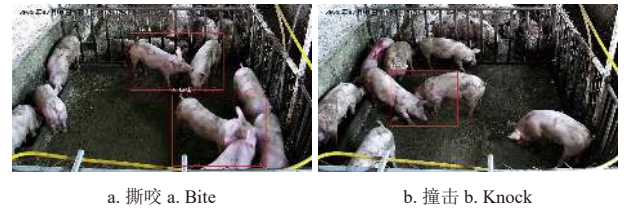


图 1 猪只攻击行为标签

Fig.1 Pig attack behavior label

2 猪只攻击行为识别算法

2.1 猪只攻击行为检测算法改进

2.1.1 改进的 YOLOX 模型

本试验中, 针对猪只攻击行为变化剧烈、图像模糊、检测精度低等问题, 提出改进的 YOLOX 模型对猪只攻击行为进行检测。

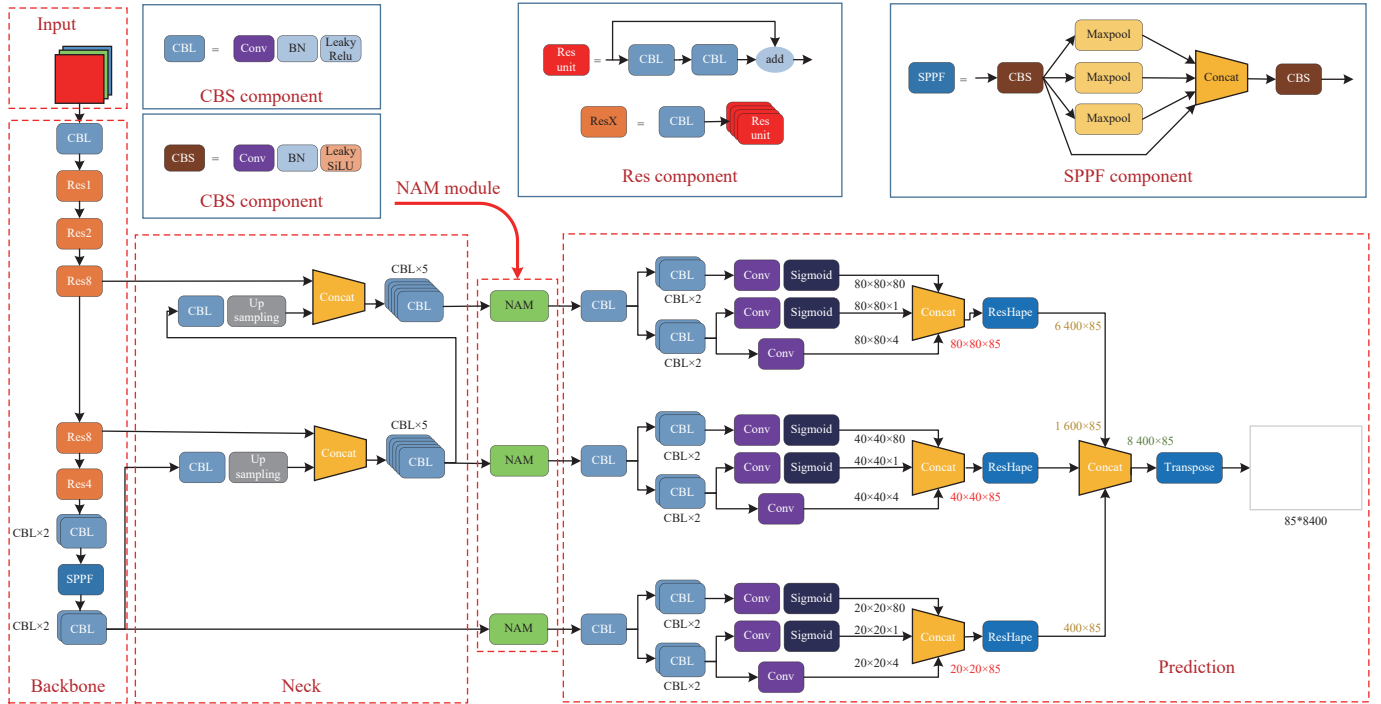
首先, 在 YOLOX 颈部末端插入归一化注意力模块^[19] (normalization-based attention module, NAM), 减少 YOLOX 在训练过程中对不显著特征的关注。其次, 将 YOLOX 中的损失函数交并比^[20] (intersection over union, IoU) 替换为广义的交并比^[21] (generalized intersection over union, GIoU), GIoU 对重叠区域和非重叠区域都进行了关注, 能更好的反映预测框和真实框的重合度, 提升 YOLOX 的检测精度。最后, 将空间金字塔池化结构 (spatial pyramid pooling, SPP) 替换为快速空间金字塔池化结构 (spatial pyramid pooling - fast, SPPF), SPPF 对数据进行了两次特征提取, 进一步过滤不显著的特征信息, 使得模型的参数量变小, 减少由于添加注意力模块导致模型参数量增大的影响。综上, 本文将改进后的模型命名为 YOLOX-NGS, 其中 N 表示 NAM, G 表示 GIoU, S 表示 SPPF。如图 2 所示。

2.1.2 NAM 注意力模块

NAM 是一种轻量级的高效注意力模块, 包括空间注意力和通道注意力两个子模块, 使用批量归一化 (batch normalization, BN) 中的比例因子测量通道的方差表明输入特征图通道的重要性, 通过抑制不太显著的变化特征, 加强通道变化特征显著特征图的重要性。如图 3 所示, 经过 5 次 CBL 层的特征提取, 对输出的特征图, 沿着空间注意力模块和通道注意力模块进行单独计算并融合, 将融合特征图乘以输入特征图进行自适应的特征细化, 突出目标的特征表达能力。

2.1.3 SPPF 结构

在 YOLO 系列中, SPP 的主要作用是实现局部特征和全局特征的特征图融合, 存在计算量大、速度慢等问题。SPPF 是在 SPP 的最大池化层前、后分别加入 CBS 模块, 对数据特征进行了 2 次提取, 如图 4 所示。SPPF 结构进一步过滤不显著特征信息, 有效减少模型参数量。



注：Conv 代表卷积块；BN 代表批量归一化；Leaky Relu 代表 Leaky Relu 激活函数；Silu 代表 Silu 激活函数；Sigmoid 代表 Sigmoid 激活函数；Maxpool 代表最大池化层；Up Sampling 代表上采样；NAM 代表归一化注意力模块；ResX 代表该模块由 X 个残差块组成，例如，Res4 代表由 4 个残差块组成的模块。Note: Conv represents convolutional blocks; BN represents batch normalization; Leaky Relu represents the Relu activation function; Silu represents the Silu activation function; Sigmoid represents the Sigmoid activation function; Maxpool represents the maximum pooling layer; Up Sampling represents up sampling; NAM stands for normalized attention module; ResX represents a module composed of X residual blocks, for example, Res4 represents a module composed of four residual blocks.

图 2 YOLOX-NGS 的网络结构

Fig.2 Network structure of YOLOX-NGS

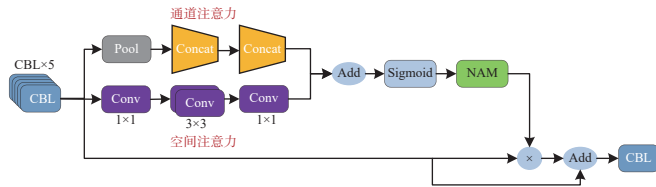


图 3 NAM 注意力模块

Fig.3 Normalization attention module

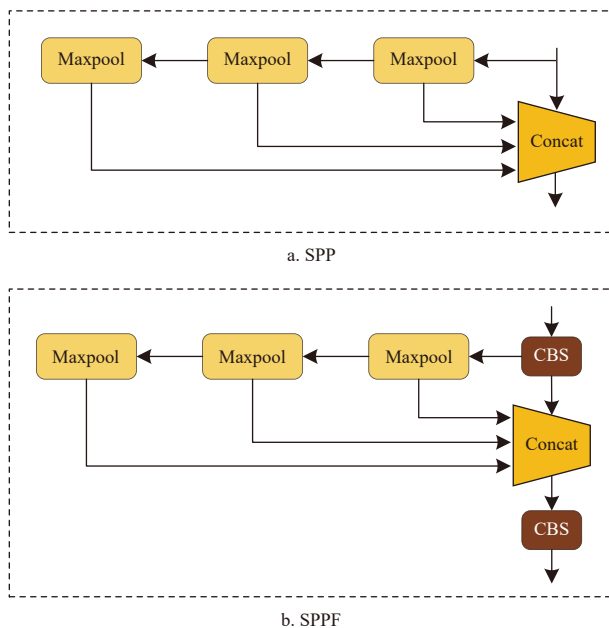


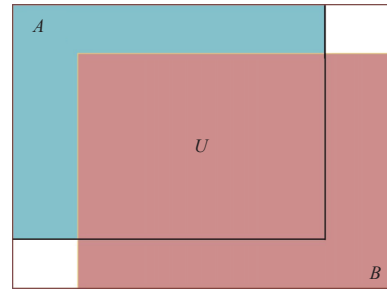
图 4 SPP 和 SPPF 结构

Fig.4 SPP and SPPF structure

2.1.4 GIoU Loss 函数

IoU 是目标检测中的一个重要指标，通过计算预测框 A 和真实框 B 的交并比，评价边界框的优劣，IoU 损失通过 IoU 计算梯度进行回归并判断目标检测结果的正确性。

GIoU 损失在 IoU 的基础上增加惩罚项， A_c 为预测框和真实框的最小外接矩形，如图 5 所示，GIoU 关注重叠区域和其他的非重叠区域，解决检测框和真实框没有相交时损失为零的问题。



注：A 代表预测边框；B 代表真实边框；U 代表 A 和 B 的并集。

Note: A represents the predicted border; B represents the true border; U represents the union of A and B.

图 5 预测框与真实框并集图

Fig.5 Union diagram of predicted box and real box

2.2 P_{AA} 和 P_{AB} 指标

本试验通过引入攻击活动比例 (P_{AA}) 和攻击行为比例 (P_{AB}) 两个指标，进一步确定攻击行为是否真实发生。试验视频段传输帧率为 30 帧/s，将第 1 次检测出攻击行

为的帧设置为首帧, 对其后续 30 帧进行猪只攻击行为的判别。当模型预测置信度大于设置的 P_{AA} 阈值时, 判别该帧为疑似攻击行为帧, 当后续帧全部判断完成后, 用疑似攻击发生帧数/总帧数 (30 帧), 得到攻击行为比例 P_{AB} 。当猪只攻击行为识别中连续帧内 P_{AB} 达到所设的阈值时, 判定该连续帧存在真实攻击行为。

$$P_{AA} = \frac{A_m}{A_t} \times 100\% \quad (2)$$

式中 A_m 为攻击行为锚框内的猪只攻击行为正确检测总数, A_t 为攻击行为检测总数。

当 P_{AA} 达到预设的阈值时, 判定该帧为疑似攻击行为帧。 P_{AB} 计算见式 (3)。式中 A_n 为疑似攻击行为帧数, S_n 为总帧数。设 $A_n=0$, 当某一帧内的攻击行为置信度大于所设的 P_{AA} 阈值时, $A_n=A_n+1$, 当 P_{AB} 达到预设的阈值时, 判定连续帧内存在攻击行为。

$$P_{AB} = \frac{A_n}{S_n} \times 100\% \quad (3)$$

2.3 模型训练与评价指标

试验环境操作系统为 Windows 10, 深度学习框架 Pytorch 1.2.0, 显卡为 NVIDIA GeForce RTX 3 060 显卡, 显存 12GB。输入图片分辨率为 640×640, 模型训练 100 个周期, 采用 Adam 优化器。

本研究对猪只攻击行为检测模型的评价指标采用深度学习目标检测常用评价指标, 包括准确率、召回率、F1 值、平均精度 (average precision, AP) 和精确率^[13-16] (A_{cc})。 A_{cc} 在本研究中表示判断猪只攻击行为正确次数与总数的比例。

3 结果与分析

3.1 猪只攻击行为检测结果与分析

3.1.1 模型对比试验

本试验选择经典的单阶段 SSD^[22] 模型、针对小目标的 CenterNet^[23] 模型、YOLOv4^[24] 模型和 YOLOv5^[25] 模型与 YOLOX^[26] 模型进行对比。对比试验结果见表 2。

如表 2 所示, YOLOX 模型的平均精度为 90.77%, 相较于其他模型提高 1.11~8.15 个百分点, 在猪只攻击行为检测中具有更优的性能。

表 2 不同模型的试验结果对比

Table 2 Comparison of experimental results with different models

模型 Model	准确率 Precision	召回率 Recall	F1 值 F1-Score	平均精度 Average precision(AP)	%
SSD	65.55	75.73	70.27	82.62	
CenterNet	82.29	76.70	79.40	85.48	
YOLOv4	85.87	76.70	81.03	88.54	
YOLOv5	91.11	79.61	84.97	89.66	
YOLOX	86.46	80.58	83.42	90.77	

3.1.2 注意力模块对模型性能的影响

本试验在颈部末端加入不同注意力模块进行对比试验, 对比试验结果见表 3。

表 3 不同注意力模块的试验结果

Table 3 The test result of different attention module

模型 Model	准确率 Precision	召回率 Recall	F1 值 F1-Score	平均精度 AP	%
YOLOX	86.46	80.58	83.42	90.77	
YOLOX-SE	81.08	87.38	84.11	90.61	
YOLOX-BAM	85.44	85.44	85.44	91.00	
YOLOX-CBAM	89.36	81.55	85.28	91.12	
YOLOX-NAM	87.50	88.35	87.92	93.27	

如表 3 所示, YOLOX-SE、YOLOX-BAM、YOLOX-CBAM 模型相对于 YOLOX 模型平均精度最高提升 0.35 个百分点, F1 值最高提升 2.02 百分点; YOLOX-NAM 模型相对于 YOLOX 模型平均精度提升 2.50 个百分点, F1 值提升 4.50 个百分点。集成 NAM 注意力模块有效抑制不太突出的特征, 增强变化明显的区域特征信息; 融合空间注意力和通道注意力的 NAM 对群养猪只攻击行为的识别更为精确, 能更好解决因猪只攻击行为动作剧烈、目标区域模糊导致识别精度下降的问题。

如图 6 所示, YOLOX 模型对猪只攻击行为检测时出现置信度低、漏检、误检和预测框偏移的问题。本试验中, 集成 NAM 注意力模块, 加强群养猪攻击行为特征的表达能力, 提高目标行为的预测置信度 (图 6a); 同时, 提升不同局部信息的捕获能力, 有效减少漏检、误检情况 (图 6b~6c); 对预测框偏移的问题, 集成 NAM 注意力模块尚未解决 (图 6d)。

3.1.3 损失函数选择对模型性能的影响

本试验选择不同损失函数进行对比试验, 对比试验详见表 4。

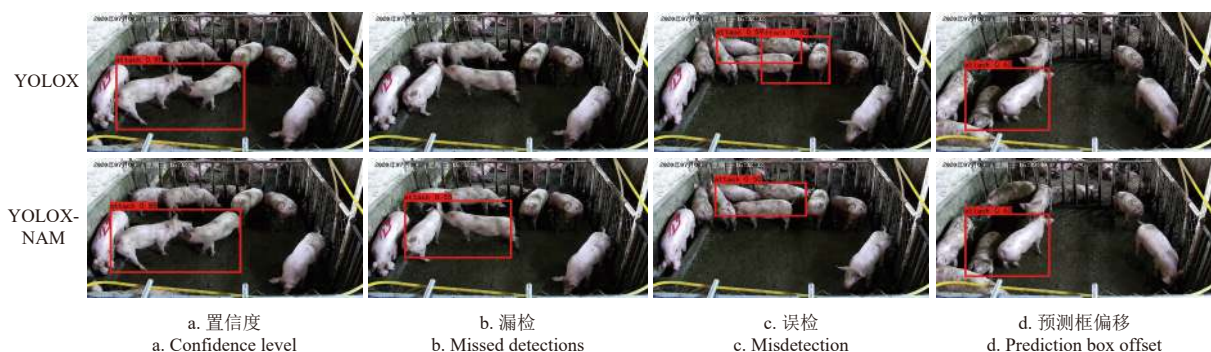


图 6 YOLOX 和 YOLOX-NAM 检测结果比较

Fig.6 Comparison of detection results of YOLOX and YOLOX-NAM

表 4 不同损失函数的试验结果

Table 4 The test result of different IoU Loss %				
模型 Model	准确率 Precision	召回率 Recall	F1 值 F1-Score	平均精度 AP
YOLOX	86.46	80.58	83.42	90.77
YOLOX-CIoU	84.26	88.35	86.26	91.07
YOLOX-DIoU	87.88	84.47	86.14	92.00
YOLOX-EIoU	83.19	91.26	87.04	91.98
YOLOX-GIoU	84.91	87.38	86.13	92.89

如表 4 所示，YOLOX-GIoU 拥有最高的平均精度。YOLOX-GIoU 模型相较于 YOLOX 模型的平均精度提升 2.12 个百分点。在本试验中，过多的关注边界框和中心

点距离等问题会导致猪只攻击行为检测精度低，因此，本试验将 IoU 损失函数替换为 Giou 损失函数，以提升模型的准确率。

图 7 为 YOLOX 和 YOLOX-GIoU 对猪只攻击行为的检测效果，Giou 损失函数能更好的反映预测框和真实框的重合度，提高猪只攻击行为的预测置信度（图 7a）。IoU 和 Giou 仅通过计算预测框和真实框交并比评价边界框的优劣，均出现预测框缺失（图 7b）。Giou 损失引入惩罚项，在不重叠的情况下，预测框会向真实框移动，解决猪体黏连和遮挡导致的误检，并且有效修正预测框偏移情况（图 7c~7d）。

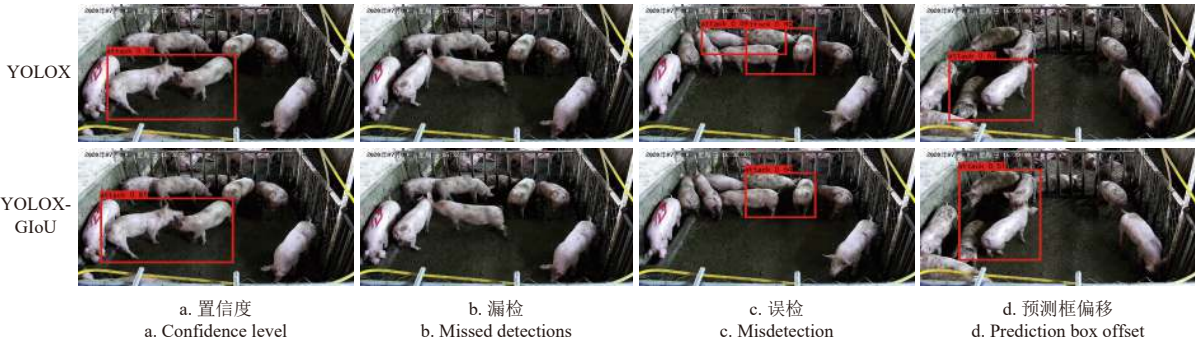


图 7 YOLOX 和 YOLOX-GIoU 检测结果比较

Fig.7 Comparison of detection results of YOLOX and YOLOX-GIoU

3.1.4 消融试验

为明确不同改进点对模型整体性能的影响，验证各改进点的有效性，本文设计消融试验从注意力模块、损失函数、特征金字塔多方面测试改进方法的有效性，结果如表 5 所示。

表 5 消融试验结果

Table 5 Results of ablation experiment								
模型 Model	改进策略 Improvement strategy			准确率 Precision/	召回率 Recall/	F1 值 F1-Score/	平均精度 AP/%	参数量 Parameters/
	NAM	SPPF	GIoU	%	%	%		MB
YOLOX	-	-	-	86.46	80.58	83.42	90.77	34.35
	√	-	-	87.50	88.35	87.92	93.27	34.76
	-	√	-	88.00	85.44	86.70	91.75	34.22
	-	-	√	84.91	87.38	86.13	92.89	34.33
	√	√	-	81.90	92.23	86.76	93.77	34.57
	√	-	√	88.07	93.20	90.56	94.24	34.74
	-	√	√	80.00	93.20	86.10	93.53	34.29
	√	√	√	95.05	93.20	94.12	97.57	34.58

由表 5 结合表 2 可知，YOLOX-NGS 与 SSD、CenterNet、YOLOv4、YOLOv5 相比，平均精度分别提

高了 14.95、12.09、9.03、7.91 个百分点。YOLOX-NGS 与 YOLOX-NAM、YOLOX-SPPF、YOLOX-GIoU 相比，平均精度分别提高了 4.30、5.82、4.68 个百分点。各个改进点对模型的检测性能都有不同的提升，在原骨干网络上分别集成 NAM 模块、替换 Giou 损失函数和 SPPF 特征金字塔结构，与原 YOLOX 模型相比的平均精度分别提升 2.5、2.12 和 0.98 个百分点。替换 SPPF 特征金字塔结构模型的参数量减少 0.1 MB，精度提升 0.98 个百分点，使集成 NAM 模块后模型参数量增大的影响降到最低。改进后模型平均精度从 90.77% 提升到 97.57%，提升 6.8 个百分点；模型的参数量从最高的 34.7 MB 降低到 34.5 MB，降低 0.2 MB，这表明改进后的 YOLOX 模型平均检测精度提升较大、参数量变化较小。综上所述，本试验中 YOLOX-NGS 模型在整体性能上超越所有模型。

如图 8 所示，YOLOX-NGS 能够准确、快速的检测出猪只攻击行为，提高猪只攻击行为检测置信度（图 8a），同时 YOLOX-NGS 解决了 YOLOX 中猪只攻击行为漏检、误检和预测框偏移问题（图 8b~8d）。

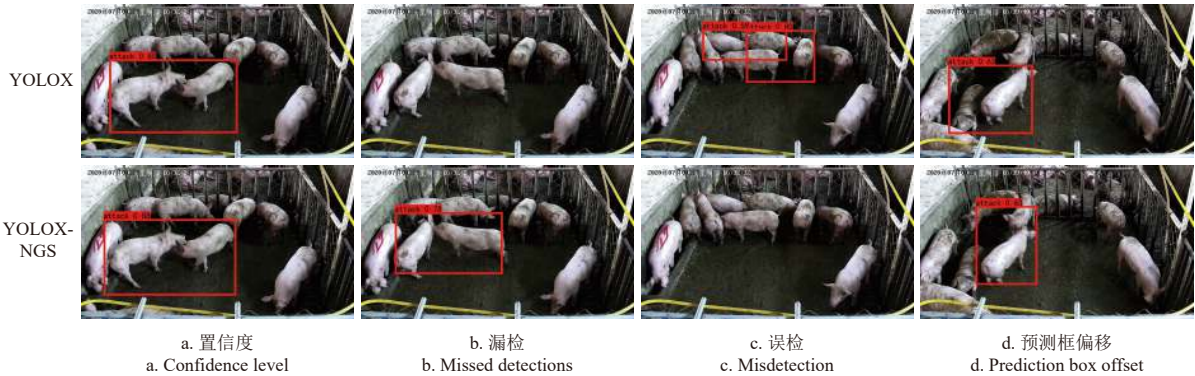


图 8 YOLOX 和 YOLOX-NGS 检测结果比较

Fig.8 Comparison of detection results of YOLOX and YOLOX-NGS

3.2 猪只攻击行为识别结果与分析

为验证猪只攻击行为识别优化策略的有效性,选取攻击频繁的视频段进行测试和评价。本试验引入的 P_{AA} 和 P_{AB} 指标对猪只攻击行为识别准确率有较大的影响, P_{AA} 阈值过小,容易把未发生攻击行为的帧误判为发生攻击行为, P_{AA} 阈值过大,容易把发生攻击行为的帧误判为未发生攻击行为。

本试验对存在攻击频繁的猪只日常行为视频段进行测试,基于不同 P_{AA} 和 P_{AB} 阈值组合的测试结果如表 6 所示。

表 6 不同攻击活动比例和攻击行为比例阈值组合下的猪只攻击行为识别结果

Table 6 Results of attack behavior of pigs under different proportion of attack activities (P_{AA}) and proportion of attack behavior (P_{AB}) threshold combinations

P_{AA} 阈值 P_{AA} threshold	P_{AB} 阈值 P_{AB} threshold	精准率 Accuracy/%	准确率 Precision/%	召回率 Recall/%
0.1	0.3	84.99	87.00	84.47
	0.4	85.08	87.13	85.44
	0.5	92.46	97.85	88.35
0.2	0.3	97.57	95.05	93.20
	0.4	98.55	95.00	92.23
	0.5	96.76	95.96	92.23
0.3	0.3	95.32	96.94	92.23
	0.4	92.46	97.85	88.35
	0.5	90.50	95.50	85.00

1) 当 P_{AA} 设置过大 (>0.3) 或者过小 (<0.1) 都会导致不同程度的漏检,因此本试验选取 P_{AA} 阈值应介于 $0.1\sim0.3$ 。当 P_{AA} 阈值为 0.2 时,精准率、准确率和召回率的值都较高, P_{AA} 上下浮动 0.1 时,虽然准确率有所提高,但召回率明显下降。

2) 当 P_{AB} 设置过大 (>0.5) 或者过小 (<0.3) 时,识别精确率有较大下降,因此本试验选取的 P_{AB} 阈值应介于 $0.3\sim0.5$ 。当 P_{AA} 阈值为 0.2、 P_{AB} 阈值为 0.4 时,攻击行为识别效果最佳,精准率、准确率和召回率分别达到 98.55%、95.00% 和 92.23%。

本试验采用 YOLOX-NGS 算法的准确率为 97.57%,达到实用化的标准,有助于推进猪只攻击行为识别研究,引入攻击活动比例 (P_{AA}) 和攻击行为比例 (P_{AB}) 提升猪只攻击行为识别的有效性,为猪只攻击行为识别实用化提供一种可靠的技术方案。

4 结 论

本试验在检测网络结构中进行归一化注意力模块 (normalization-based attention module, NAM) 添加、GIoU 损失函数替换和 SPPF 金字塔结构优化,来完成不同环境和不同影响条件下的单帧猪只攻击行为目标检测。引入攻击活动比例 (proportion of attack activities, P_{AA}) 和攻击行为比例 (proportion of attack behavior, P_{AB}) 提升猪只攻击行为识别的有效性。主要结论如下:

1) 本文提出的 YOLOX-NGS 深度学习检测模型平均精度为 97.57%,模型大小为 34.58 MB。与 SSD、CenterNet、YOLOv4、YOLOv5 相比,平均精度分别提高了 14.95、12.09、9.03、7.91 个百分点;与改进策略

YOLOX-NAM、YOLOX-SPPF、YOLOX-GIoU 相比,平均精度分别提高了 4.30、5.82、4.68 个百分点。综上,YOLOX-NGS 对猪只攻击行为检测优势明显,与其他检测算法和改进策略相比,YOLOX-NGS 平均精度提升 $4.30\sim14.95$ 个百分点,模型参数量和检测速率具有一定的优势,有助于猪只攻击行为在线监测在实际养殖场景中的应用。

2) 本文提出猪只攻击行为识别 P_{AA} 和 P_{AB} 指标。 P_{AA} 和 P_{AB} 阈值分别设置为 0.2 和 0.4 时,群养猪只的攻击行为识别效果最佳,识别的平均精度高达 98.55%。满足实际应用场景下猪只攻击行为识别的精度要求,可为群养猪只攻击行为识别提供技术支持。

[参 考 文 献]

- [1] 叶建伟,薛亚辉,王明宇,等.行为自动检测技术在生猪健康和福利方面的研究进展[J].家畜生态学报,2019,40(10): 8-13, 7.
YE Jianwei, XUE Yahui, WANG Mingyu, et al. Research progress on automatic pig behavior detection technology in animal health and welfare[J]. Acta Ecologiae Animalis Domastici, 2019, 40(10): 8-13, 7. (in Chinese with English abstract)
- [2] 贺稚非,余思洁,李洪军,等.非洲猪瘟病毒生物学特性及在猪肉生产中的防控措施研究[J].农业工程学报,2019,35(10): 299-304.
HE Zhifei, YU Sijie, LI Hongjun, et al. Biological characteristics of African swine fever virus and its prevention and control measures in pork food production[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(10): 299-304. (in Chinese with English abstract)
- [3] VIAZZI S, ISMAYILOVA G, OCZAK M, et al. Image feature extraction for classification of aggressive interactions among pigs[J]. Computers and Electronics in Agriculture, 2014, 104: 57-62.
- [4] OCZAK M, VIAZZI S, ISMAYILOVA G, et al. Classification of aggressive behaviour in pigs by activity index and multilayer feed forward neural network[J]. Biosystems Engineering, 2014, 119(4): 89-97.
- [5] JONGUK L, LONG J, DAIHEE P, et al. Automatic recognition of aggressive behavior in pigs using a kinect depth sensor[J]. Sensors, 2016, 16(5): 631.
- [6] CHEN C, ZHU W, MA C, et al. Image motion feature extraction for recognition of aggressive behaviors among group-housed pigs[J]. Computers and Electronics in Agriculture, 2017, 142: 380-387.
- [7] 肖德琴,曾瑞麟,周敏,等.基于 DH-YoloX 的群养马岗鹅关键行为监测[J].农业工程学报,2023,39(2): 142-149.
XIAO Deqin, ZENG Ruilin, ZHOU Min, et al. Monitoring the vital behavior of Magang geese raised in flocks based on DH-YoloX[J]. Transactions of the Chinese Society of Agricultural

- Engineering (Transactions of the CSAE), 2023, 39(2): 142-149. (in Chinese with English abstract)
- [8] 黄志杰, 徐爱俊, 周素茵, 等. 融合重参数化和注意力机制的猪脸关键点检测方法[J]. 农业工程学报, 2023, 39(12): 141-149.
- HUANG Zhijie, XU Aijun, ZHOU Suyin, et al. Key point detection method for pig face fusing reparameterization and attention mechanisms[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(12): 141-149. (in Chinese with English abstract)
- [9] 齐咏生, 焦杰, 鲍腾飞, 等. 基于自适应注意力机制的复杂场景下牛脸检测算法[J]. 农业工程学报, 2023, 39(14): 173-183.
- QI Yongsheng, JIAO Jie, BAO Tengfei, et al. Cattle face detection algorithm in complex scenes using adaptive attention mechanism[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(14): 173-183. (in Chinese with English abstract)
- [10] YANG Q M, XIAO D Q, LIN S C. Feeding behavior recognition for group-housed pigs with the Faster R-CNN[J]. Computers and Electronics in Agriculture, 2018, 144: 453-460.
- [11] YANG A, HUANG H, ZHENG C, et al. High-accuracy image segmentation for lactating sows using a fully convolutional network[J]. Biosystems Engineering, 2018, 76: 36-47.
- [12] 杨秋妹, 肖德琴, 张根兴. 猪只饮水行为机器视觉自动识别[J]. 农业机械学报, 2018, 49(6): 232-238.
- YANG Qiumei, XIAO Deqin, ZHANG Genxing. Automatic pig drinking behavior recognition with machine vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(6): 232-238. (in Chinese with English abstract)
- [13] 李菊霞, 李艳文, 牛帆, 等. 基于 YOLOv4 的猪只饮食行为检测方法[J]. 农业机械学报, 2021, 52(3): 251-256.
- LI Juxia, LI Yanwen, NIU Fan, et al. A method for detecting dietary behavior in pigs based on YOLOv4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(3): 251-256. (in Chinese with English abstract)
- [14] 高云, 陈斌, 廖慧敏, 等. 群养猪侵略性行为的深度学习识别方法[J]. 农业工程学报, 2019, 35(23): 192-200.
- GAO Yun, CHEN Bin, LIAO Huimin, et al. Recognition method for aggressive behavior of group pigs based on deep learning[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(23): 192-200. (in Chinese with English abstract)
- [15] ZHENG C, ZHU X M, YANG X F, et al. Automatic recognition of lactating sow postures from depth images by deep learning detector[J]. Computers and Electronics in Agriculture, 2018, 147: 51-63.
- [16] 薛月菊, 朱勋沐, 郑婵, 等. 基于改进 Faster R-CNN 识别深度视频图像哺乳母猪姿态[J]. 农业工程学报, 2018, 34(9): 189-196.
- XUE Yueju, ZHU Xunmu, ZHENG Chan, et al. Lactating sow postures recognition from depth image of videos based on improved Faster R-CNN[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2018, 34(9): 189-196. (in Chinese with English abstract)
- [17] 房俊龙, 胡宇航, 戴百生, 等. 采用改进 CenterNet 模型检测群养生猪目标[J]. 农业工程学报, 2021, 37(16): 136-144.
- FANG Junlong, HU Yuhang, DAI Baisheng, et al. Detection of group-housed pigs based on improved CenterNet model[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(16): 136-144. (in Chinese with English abstract)
- [18] 曾繁国, 朱君, 王海峰, 等. 改进帧间差分-深度学习识别群养猪只典型行为[J]. 农业工程学报, 2022, 38(15): 170-178.
- ZENG Fanguo, ZHU Jun, WANG Haifeng, et al. Typical behavior recognition of herd pigs based on improved frame difference and deep learning[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(15): 170-178. (in Chinese with English abstract)
- [19] LIU Y, SHAO Z, TENG Y, et al. NAM: Normalization-based attention module[EB/OL]. (2021-11-24)[2022-12-15]. <https://arxiv.org/abs/2111.12419>.
- [20] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2018, 42(2): 318-327.
- [21] HAMID R, NATHAN T, JUNYOUNG G, et al. Generalized intersection over union: A metric and a loss for bounding box[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, April 1-9, 2019, Los Angeles, America: IEEE, 2019: 658-666.
- [22] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector[C]//European Conference on Computer Vision, October 8-16, 2016, University of Amsterdam, Amsterdam: ECCV, 2016: 21-37.
- [23] DUAN K, BAI S, XIE L, et al. Centernet: Keypoint triplets for object detection[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, April 1-9, 2019, Los Angeles, America: IEEE, 2019: 6569-6578.
- [24] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 14-19, 2020, Seattle, USA: IEEE, 2020: 1-17.
- [25] ZHANG K, WANG C, YU X, et al. Research on mine vehicle tracking and detection technology based on YOLOv5[J]. Systems Science & Control Engineering, 2022, 10(1): 347-366.
- [26] GE Z, LIU S, WANG F, et al. YoloX: Exceeding yolo series in 2021[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 19-25, 2021, Virtual: IEEE, 2021: 1-7.

Recognizing attack behavior of herd pigs using improved YOLOX

LI Yanwen, LI Juxia^{*}, NA Tengxiao, ZHI Qingyu, DUAN Lei, ZHANG Pengpeng

(College of Information Science and Engineering, Shanxi Agricultural University, Taigu 030801, China)

Abstract: Image data was collected at the Pig Breeding Base in Fenxi County, Linfen City, Shanxi Province in July 2020. Nine 5-month-old fattening pigs were selected to raise in a closed pig house. Hikvision DS-2CD3345D-I model camera was used in a downward tilt angle of 60 degrees to collect data under incandescent light. This angle was utilized to obtain the rich behavioral features of pigs, in order to avoid large-scale occlusion, compared with the head-up and overhead views. In the process of data collection, the daily behavior videos of pigs were first repeatedly observed, and 185 video clips of pigs with aggressive behavior were extracted; Inter frame difference method was used to extract the key frames from these video clips. Slow pig movement and long rest time were removed as well. An improved YOLOX model was proposed to identify the typical attack behaviors of herd pigs, such as impact, ear biting, and tail biting. The high accuracy and effectiveness were achieved to reduce pig stacking and adhesion in complex pen environments. Firstly, a Normalization based Attention Module (NAM) was added to obtain the global information about the YOLOX neck; Secondly, the loss function IoU Loss in the YOLOX was replaced with the GIoU to improve the recognition accuracy; Finally, the real-time performance of the model was realized to enhance feature extraction and detection efficiency. Feature pyramid structure SPP was lightweight to SPPF. The experiment showed that the integrated NAM modules, GIoU Loss replacing, and SPPF feature pyramid structures in the original backbone network improved the average accuracy of the model by 2.50, 2.12, and 0.98 percentage points, respectively. The model with SPP feature pyramid structure reduced the parameter by 0.1 MB and improved the accuracy by 0.98 percentage points, indicating the minimum impact of the model parameter after the integrated NAM module. The average accuracy of the improved model increased from 90.77% to 97.57%, with an increase of 6.80 percentage points; The parameter quantity decreased from the highest 34.7 to 34.5 MB with a decrease of 0.2 MB. In addition, there was the continuous attack behavior of pigs in the low credibility of single-frame images. Two optimization indicators (proportion of attack activities (P_{AA}) and proportion of attack behavior (P_{AB})) were introduced to further confirm whether the attack behavior occurred. When the P_{AA} and P_{AB} thresholds were 0.2 and 0.4, respectively, the recognition accuracy (Accuracy) reached 98.55%. Video segments with frequent attacks were selected to verify the effectiveness of the optimization. Usually, P_{AA} and P_{AB} posed a significant impact on the recognition of pig aggressive behavior; If the threshold set was too small, it was easy to misjudge frames without attack behavior as having occurred; If the threshold set was too large, the frame without the attack was assumed as the occurrence. The experimental results show that the improved YOLOX model was achieved in the high-precision recognition of pig attack behavior by the integrated P_{AA} and P_{AB} . The finding can provide effective reference and technical support for the intelligent monitoring of herd health pigs.

Keywords: pig; behavior recognition; detection; attack behavior; attention mechanism; loss function; feature pyramid